Lect15 - 2020-03-31 - handwritten notes

$$\text{Sigmoid} = \frac{1}{1+e^{-z}}$$

$$\hat{y}_i = \sigma(\theta_0^{(3)} + \theta^{(3)}\phi(\theta_0^{(2)} + \theta^{(2)}\phi(\theta_0^{(1)} + \theta^{(1)} x_i)))$$

$$\Uparrow$$

$$\mathcal{L}(\hat{y}_i, y_i)$$

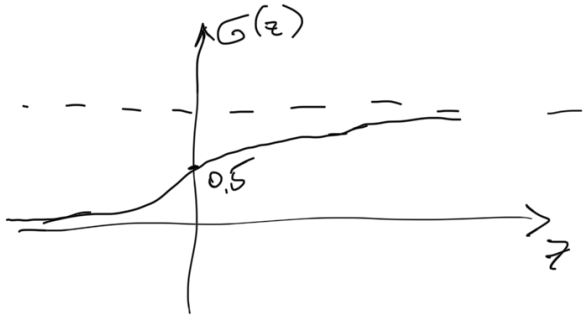
$$z^{(1)} = x_i \theta^{(1)}$$
$$a^{(1)} = \phi(z^{(1)})$$
$$z^{(2)} = a^{(1)} \theta^{(2)}$$
$$a^{(2)} = \phi(z^{(2)})$$
$$z^{(3)} = a^{(2)} \theta^{(3)}$$
$$\hat{y} = \sigma(z^{(3)})$$

$$\frac{\partial \mathcal{L}}{\partial z} \quad \frac{\partial \hat{y}}{\partial z} \quad \frac{\partial z^{(3)}}{\partial \theta^{(3)}}$$

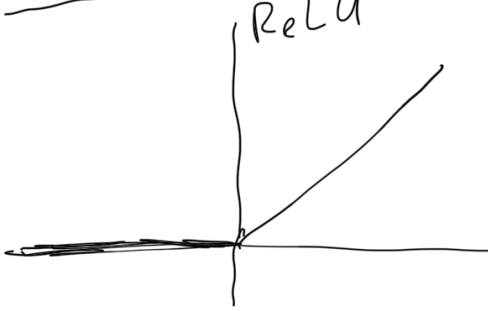
$$\frac{\partial z^{(3)}}{\partial a^{(2)}} \quad \frac{\partial a^{(2)}}{\partial z^{(2)}} \quad \frac{\partial z^{(2)}}{\partial \theta^{(2)}}$$

$$\frac{\partial \mathcal{L}}{\partial \theta^{(1)}}$$

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

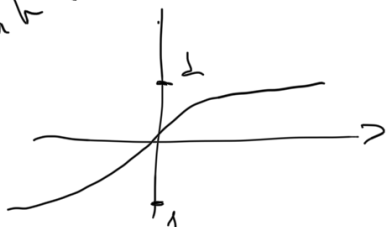


$$\text{ReLU}(z) = \max(0, z)$$



$$\exp(z)$$

$$\tanh(z)$$



$$\frac{\partial \text{ReLU}(z)}{\partial z} =$$

$X \sim N \times f$ N - кол-во объектов $= 3000$
f - кол-во признаков $= 2$

$X \to X^*$

(H) $\theta^{(1)*}$ 18шт $\theta^{(2)*}$ 28 $\theta^{(3)*}$ 15

$\theta^{(1)*} \sim 3 \times 6$ — 18
$\theta^{(2)*} \sim 7 \times 4$, — 28
$\theta^{(3)*} \sim 5 \times 3$ — 15

$18 \to 3 \times 6$
$28 \to 7 \times 4$
$5 \times 3 \to 15$

$N \times (f+1) \cdot (f+1) \times$
$3000 \quad 3 \quad 3$

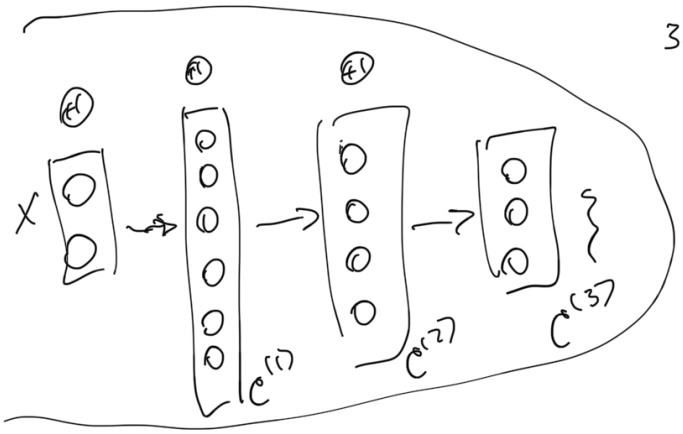
$a^{(1)*} \sim N \times (c^{(1)}+1)$

$z^{(2)} = a^{(1)*} \cdot \theta^{(2)*}$

$N \times (c^{(1)}+1) \cdot (c^{(1)}+1) \cdot c$
$3000 \quad 7 \quad 7 \quad 4$

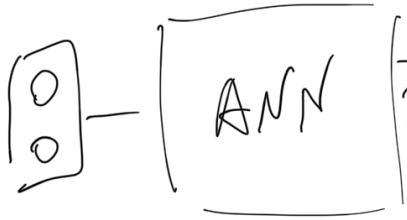
$z^{(3)} = a^{(2)*} \cdot \theta^{(3)*}$

$N \times (c^{(2)}+1) \cdot (c^{(2)}+1) \times c$
$3000 \quad 5 \quad 5 \quad 3$



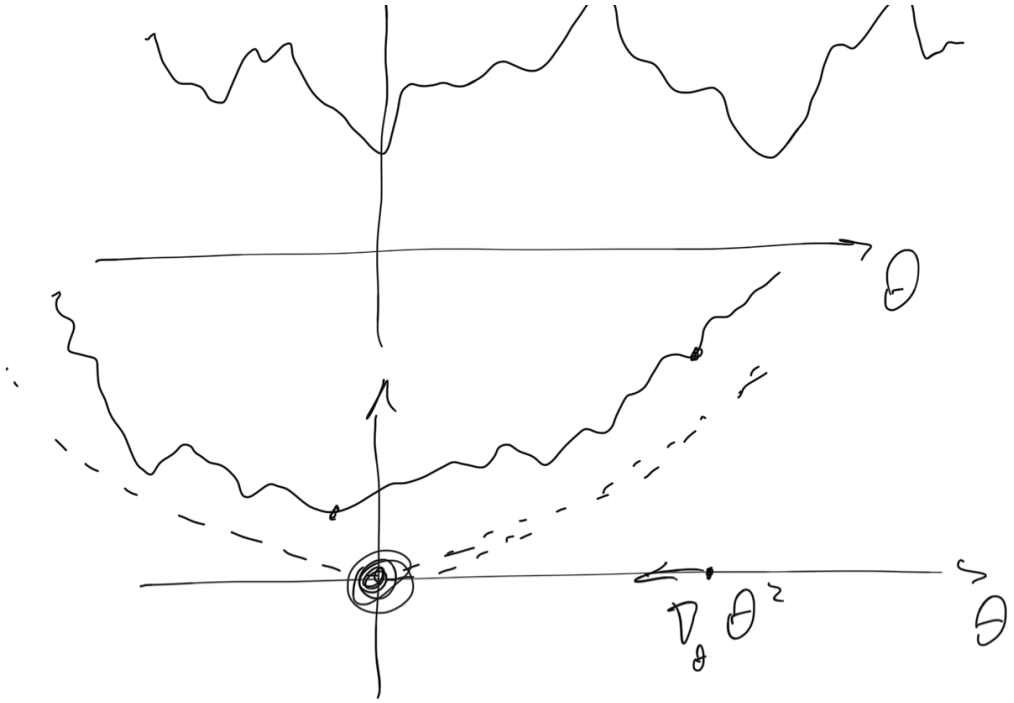
Sequential

$\|\theta\| = (f+1)c^{(1)} +$
$+ (c^{(1)}+1) \cdot c^{(2)}$
$+ (c^{(2)}+1) \cdot c^{(3)}$



$\mathcal{L}(p(x,\theta), y) + \|\theta\|$

$\nabla_\theta \mathcal{L}$


θ
$\nabla_\theta \theta^2$
θ